

Note: Line numbers listed in the response correspond to the submitted Word document containing the revised manuscript.

Reviewer #1

Nitrate contamination in irrigated stream-aquifer systems is a serious problem in agricultural watershed. Numerical modelling and relevant sensitivity analysis are important methods for understanding of nitrogen fate and transport, as well as making remediation strategies. This study used a nitrogen fate and transport groundwater model and the revised Morris sensitivity analysis method to identify the spatially-varying influence of system factors on nitrate fate and transport in a regional-scale irrigated hydro-agricultural system. Some results were valuable for future data collection and remediation strategies in the study area. On the whole, the paper was well written. Some minor improvements and corrections are needed.

1. According to the title, the spatially-distributed influences should be the emphasis in this paper. But in the abstract there are no such descriptions or conclusions. some important conclusions should be added in the abstract.

Response: The abstract has been modified to include more details regarding the spatial variability of factor dominance. The modified text read:

Lines 31-38:

“Results suggest that fertilizer loading, crop uptake, and heterotrophic denitrification govern NO_3 fate and transport for the majority of the study area, although their order of influence on NO_3 groundwater concentration and mass leaching varies according to crop type and command area. Canal NO_3 concentration and rates of autotrophic denitrification, nitrification, and humus decomposition also dominate or partially dominate in other locations. Each factor, with the exception of O_2 reduction rate, is the dominating influence on NO_3 groundwater concentration at one or more locations within the study area.”

2. In your model, there are 7 vertical layers. Each layer has the same depth for each grid cell, which means you don't consider the topography? Is the groundwater table keeping constant or can be changed in different seasons?

Response: Only the top 3 layers have the same thickness across the model domain. Layers 4-6 vary according to the saturated thickness of the aquifer. This is now described in the text:

Lines 277-278:

“Thickness of layers 4, 5, and 6 varies according to saturated thickness, with layer thickness ranging from 2.8 m to 12.6 m.”

The groundwater hydraulic head changes from week to week (the MODFLOW model uses weekly time steps) according to changes in infiltrating rainfall, irrigation water, pumping, and groundwater-surface interactions. The MODFLOW model used in this study is described on **Lines 266-273**.

3. There should be many input parameters in the UZF-RT3D model, why you chose such 9 factors as the target to analyze? Please give the explanation.

Response: Based on literature, the parameters selected represent the main fate and transport processes for Nitrogen species. Also, they were found to be the most dominant when 13 parameters were assessed in a previous study [Bailey, R.T., T.K. Gates, and M. Ahmadi (2014), *Simulating reactive transport of selenium coupled with nitrogen in a regional-scale irrigated groundwater system*. J. Hydrol. 515: 29-46], as referenced in the manuscript. This is now mentioned in the text:

Lines 150-152:

“Nine model factors are included in the assessment, with their overall influence on NO₃ fate and transport evidenced in a previous study in the region [Bailey et al., 2014].”

4. E is an environmental reduction factor that accounts for θ and T and acts to temper the reaction rates based on microbial activity, ———do θ and T have the same effects on microbial activities in different processes (i.e. nitrification, volatilization and denitrification)?

Response: All microbial processes are affected equally by soil temperature. However, water content affects nitrification, mineralization, and denitrification differently (this is based on Birkinshaw and Ewen, 2000, reference in the manuscript; also, more description is found in Bailey et al., 2013b, also referenced in the manuscript). This is now stated in the text:

Lines 262-265:

“...and E [-] is an environmental reduction factor that accounts for θ and T and acts to temper microbial activity rates [Birkinshaw and Ewen, 2000; Bailey et al., 2013b]. Nitrification, mineralization, and denitrification each have uniquely specified relationships between θ and microbial activity.”

5. In page 14, formula (5) , the left should be rfvol

Response: Thank you. This has been corrected.

6. in line 5 of page 24,it seems to have some words lost after ‘and.’.

The term $Canal_{NO_3}$ did not appear in the typeset PDF. We will ensure that is correct in the revised

Response: submission.

7. In line 10, line 17 of page 24, ‘CO2’ and ‘NO3’ should be ‘CO₂’ and ‘NO₃’?

Response: This again seems to be an issue with the typesetting. It appears as C_{O_2} in the submitted document.

8. In line 18 of page 24, ‘andshould’ should be ‘should’? or some words after ‘and’ are lost?

Response: The term $Canal_{NO_3}$ did not appear in the typeset PDF.

9. In line 19 of page 24, ‘withmonitored’ should be ‘with monitored’?

Response: The term $Canal_{NO_3}$ did not appear in the typeset PDF.

We thank Reviewer #1 for the helpful comments and suggestions.

Reviewer #2

Comment 1: The way the parameter sets are created for Morris' method is rather unusual. In typical applications realistic parameter ranges are chosen before the analysis and a (uniform) random sampling is performed (Pianosi and Wagener, 2015; Saltelli et al., 2008). Here, the parameters are varied locally around their "base values" that were found in Bailey et al. (2014) and sampled from a (log)Gaussian distribution (as far as I understand the sampling scheme). There is no elaboration, why this particular way is chosen and how it will affect the interpretations compared to a more typical application of Morris. Considering this modified parameter sampling the term "global" sensitivity analysis appears not adequate and I recommend changing it to "regional" or "local".

Response: Many Morris applications in the literature have used a uniform distribution. This approach is not different from approaches followed by other researchers. Please see Saltelli et al. (2008)¹'s discussion on page 6 and many other places on sampling the parameters according to their perspective distributions. They have used normal distribution in many of the examples. Nonetheless, our previous experience shows that results of the sensitivity analysis are not sensitive (or have negligible sensitivity) to the parameter distribution. Rather, selection of parameters range (lower and upper values) has a much stronger influence on the results. Also, Figure 5 (values of fertilizer loading, autotrophic denitrification first-order rate constant, concentration of nitrate in the irrigation canals) shows that the values used in the 280 simulation span the range specified for each parameter.

Comment: Page 1656, Line 28 and page 1657, Line 12- New references

Response: Thank you. They are very interesting papers. The references have been added to the literature review.

Comment: Page 1661, Line 3- Preceding paragraphs: please shorten. The Information provided in the study area description should be enough to Support the subsequent analysis but much more than that can be confusing for the reader

Response: The description of the study region has been shortened (753 words to 593 words).

Comment: Page 1661, Line 4- Paragraph below: this is part of the model description -2 please move

Response: The preliminary description of the model for the study region has been moved to the last paragraph of Section 2.2. Also, the description of the MODFLOW model was shortened, so as not to distract from the UZF-RT3D model description.

Comment: Page 1666, Line 6- The elaborations on the reaction module are too detailed for a model that was developed in a previous study. Please focus On the most relavant features and equations and move the detailed description to the appendix

Response: The model description was shortened:

- Equation (1) and (2) (chemical reactions for autotrophic reduction in the presence of shale) was removed.
- Equations (6), (7), (8), (10), and (11) also were removed

¹ Saltelli, A., Ratto, M., Andres, T., Campolongo, F., Cariboni, J., Gatelli, D., Saisana, M., and Tarantola, S.: Global Sensitivity Analysis. The Primer, John Wiley and Sons, West Sussex, UK, 2008.

As modified, it seems that there is a better balance between providing enough information about the model and leaving many of the details to other references.

Comment: Page 1667, Line 4- SD → please use "standard deviation"

Response: Recommended change was implemented in the revised manuscript.

Comment: Page 1667, Line 19-Also here for Morris' method some more focus /less text would be desirable.

Response: Morris SA method section was shortened in the revised manuscript: *"A high value of σ for the EE distribution signifies...independent of the values assigned to other parameters"* was removed.

Comment: Page 1667, Line 27- Why is this necessary here?

Page 1668, Line 2- Please provide some more details on the Savage score. How and why is it used here?

Response: As described on **Lines 316-321**, we used the Savage score to compare results from different SA experiments. As described in Equation 7, the Savage score is estimated from the rank of the parameter with an emphasis on the high-ranked parameters. The need for the Savage score is clarified in the revised manuscript:

Lines 322-326:

Response: *"For example, the highest ranked variable would have a score of $1/1 + 1/2 + 1/3 + \dots + 1/k$. The second ranked variable would have a score of $1/2 + 1/3 + \dots + 1/k$, and so on. Savage scores typically are preferred because they place higher emphasis on the agreement of the key drivers (i.e. higher ranked parameters), rather than the overall agreement."*

Comment: Page 1668, Line18- Creating parameter sets just around a base set of parameters is not really a global sensitivity analysis because not the entire parameter space is sampled. The authors Should clarify this here.

Response: The Morris method is considered a global SA method because it samples from entire parameter space. This is demonstrated in Figure 5 (values of fertilizer loading, autotrophic denitrification first-order rate constant, concentration of nitrate in the irrigation canals), which shows that the values used in the 280 simulation span the range specified for each parameter.

Comment: Page 1668, Line 19- why did the authors use a sampling Scheme that assumed a Gaussian distribution?

Response: Please refer to our response to Comment 1. As stated in the text, 6 of the 9 parameters (those that deal with chemical reaction rates) were perturbed using log values, since their distribution typically is log-normally distributed. In general, however, we have found that results of sensitivity analysis are sensitive to the parameter range (lower and upper values) rather than to the shape of the distribution.

Comment: Page 1668, Line 26- Why are these results mentioned and shown here? Aren't We in the methods Section?

Response: The results shown deal with the set-up of the SA (rather than the results of running the parameter sets through the UZF-RT3D model) and the input values to the UZF-RT3D model simulations. Therefore, we elected to place it in the Methods section in the original manuscript. Upon consideration of the reviewer comments, we still feel that Figure 5 shows the model set-up, and therefore belongs in the Methods section.

Comment: Page 1670, Line 15- Is there any information on the model performance/ realism? Are there observations or other studies that confirm the spatial pattern?

Response: The model results (average of the 280 simulations) did not vary significantly from the baseline model performance summarized in Bailey et al. [2014] (the reference in the manuscript), and therefore were not compared to field data in the original manuscript. Basic results are now included (comparison of average nitrate groundwater concentrations within canal command areas):

Lines 386-390:

“Average C_{NO_3} across all simulations for each command area are (average of observed field values are in parentheses) Highline 2.0 mg/L (3.1 mg/L); Catlin: 1.4 mg/L (6.1 mg/L); Rocky Ford: 1.5 mg/L (3.8 mg/L); Fort Lyon: 3.7 mg/L (1.6 mg/L); Holbrook: 1.9 mg/L (3.5 mg/L); and non-cultivated areas: 3.5 mg/L (4.2 mg/L). Average values correspond closely to results from the tested baseline model [Bailey et al., 2014].”

Comment: Page 1671, Line 26- There are results and conclusions but a discussion and Comparison to other studies is missing

Response: A general discussion section has been added (Section 4). Discussion pieces from the Results section (Section 3) have been moved to this new section (Section 4) and expanded to provide comparisons with other studies. We feel that this new format (and added comparisons with previous studies) is a great improvement of the manuscript, and thank the reviewer for the helpful comment.

Comment: Page 1672, Line 1- Somewhere within the methods the authors should provide an overview about their general approach. This will elaborate the purpose of each individual methods and make it more easy to understand why particular results are discussed more than others.

Response: The general approach is included in the original manuscript (**Lines 135-159**), before Section 2.1 begins. These two paragraphs provide an overview of the Methods, including what is focused on in the results.

Comment: Page 1672, Line 24- same as above: no comparison to other studies

Response: See response to previous comment. There are now comparisons to other studies in Section 4 (“Discussion of Results”).

Comment: Page 1675, Line 1- All subsections of the results section provide a description of the results but only few interpretation and some conclusions. A real discussion is missing, there is not a single reference to other studies.

Response: See response to previous comment.

For better readability the authors could think of a way to reduce the high usage of parameter acronyms. More tables and a shorter focused text description may be a way. Right now it is very tough to read the results.

Response: We agree that the usage of numerous acronyms within the text is tedious for the reader. The following edits have been made to improve the readability of the manuscript.

- Terms in the first paragraphs of the Results section are re-defined
- Parameter symbols are replaced with actual names in many instances throughout Section 3 and Section 4. The parameter symbol is used only if the word definition is used in the same paragraph.
- Column headings (definitions of the symbol) are added to Tables 3 and 4
- The discussion of results is moved to a separate, new section (Section 4), with only word definitions used.
- Word definitions are placed in Figures 9 and 10 next to each map.

All of the minor corrections in the text have been implemented in the revised manuscript. As stated in the response to Reviewer #1, some of these corrections are due to symbols (e.g. $Canal_{NO_3}$ in Section 4) in the Word document not appearing when the article PDF was created.

We thank Reviewer #2 for the helpful comments and suggestions.